

Voice Based Email for Visually Challenged

Uvaish Jafri¹ Abdul Ahad² Saurabh Srivastava³ Kartikey Mishra⁴

^{1,2,3,4}Department of Information Technology

^{1,2,3,4}Noida Institute of Engineering Technology, AKTU, Noida, India

Abstract— One of the mostly used forms of communication among the people is Email. Lot of confidential and urgent information is exchanged over emails in today's time. There are about 253 million visually impaired people worldwide. These visually impaired people are facing a problem of communication. Since, technology is growing day by day these types to visually challenged people feel that they are more challenged. Therefore, proposed a Voice based Email System using AI that will make email system very easily accessible to visually challenged people and also help society. Accessibility is the most important feature that is considered while developing this system. Any system is called accessible only if both able and disable people can use it easily.

Keywords: Electronic Mail (EMail), Speech_To_Text conversion (STT), Text_To_speech Conversion (TTS)

I. INTRODUCTION

The technologies are growing very fast day by day this has made the lifestyle of people so easy as mostly all work can be done in less amount of time with accuracy and efficiency.

Communication is one of those fields that have grown to next level with the advancement in technology and the availability of Internet. Technologies have made communication so easy that distance has become a negligible parameter in communication.

When we think of communication using internet, the first thing that we come in our mind is communication via email. Email is one of the most reliable ways for exchange of some important information and also email is used worldwide, but for accessing internet a person must be able to see. There are millions of people who are blind or visually challenged who are not able to see the screen; keyboards therefore they are not able to access the internet.

One of the major drawbacks that sets in is that accessing emails or on a whole any page on the internet requires a person to have visual capabilities. This means that a visually challenged person can in no way take the benefits of the facilities provided by the internet thus rendering the technology useless.

In this way, they are very far away from email communication and internet world. These blind people cannot use the existing email system, they cannot send, receive emails and cannot read the information shared through email, therefore the existing systems are not easily accessible to them.

To access the internet the person must be able to read what is written on the screen so, this makes internet useless technology for visually challenged people. There is only one way by which a visually challenged person can send an E-mail is, they have to tell the entire content of the mail to a third person so that the third person can compose the mail and send on the behalf of the visually challenged person.

But this approach does not take us to the solution of the problem. Every time finding a third person is not possible for a visually challenged person and also sometimes the

content can be personal, for maintaining the Integrity of the Specifications.

Therefore, for helping these people and developing society authors have come up with this idea that helps a visually challenged person by providing ability to send and receive emails throw voice commands without using any keyboard and visual thing.

Artificial Intelligence for Speech Recognition: Artificial intelligence (AI) is a technology used for creating intelligent systems and machines that simulate human intelligence. Some artificial intelligence applications include various expert systems, natural language processing (NLP), machine vision and speech recognition. To understand and analyse human language such as English by extracting metadata from keywords, emotions, relations and concepts is Natural Language Processing.

II. SPEAKER RECOGNITION

Speaker recognition has generally been viewed as a problem of verifying or recognizing a particular speaker in a segment of speech spoken by a single speaker. But for some applications of interest the problem is to verify or recognize particular speakers in a segment of speech in which multiple speakers are present. Automatic systems need to be able to segment the speech among the speakers present and/or to determine whether speech by a particular speaker is present and where in the segment this speech occurs. Speaker recognition process encompasses three terms identification, verification and diarization. In automatic speaker identification, there is no priori identity claim, the system decides who the person is, or the person is known or unknown.

III. SYSTEM ANALYSIS

A. Existing System

In previous work, blind people does not send email using the system. The multitude of email types along with the ability setting enables their use in nomadic daily contexts. But these emails are not useful in all types of people such as blind people they can't send the email. Audio based email are only preferable for blind peoples. They can easily respond to the audio instructions. In this system is very rare. So there is less chance to available this audio based email to the blind people.

B. Proposed System

We describe the voicemail system architecture that can be used by a blind person to access e-mails easily and efficiently. The contribution made by this research has enabled the blind people to send and receive voice-based e-mail messages in their native language with the help of a computer or a mobile device. Our proposed system GUI has been evaluated against the GUI of a traditional mail server. We found that our proposed architecture performs much better than that of the existing GUI'S.

IV. DESIGN

A. User Interface Design:

The user interface is designed using Adobe Dreamweaver CS3. The complete website focuses more on efficiency in understanding the IVR rather than the look and feel of the system as the system is primarily developed for the blind people to whom the look and feel won't be of that primary importance as the efficiency of understanding the prompting would be.

B. Database Design:

Our system maintains a database for user validation and storing mails of the user. There are a total of five tables. The Inbox, Sent-Mail schemas will store all mails of the respective service that belongs to that particular user.

C. System Design:

Our System is voice oriented. When user is over every legal space in website, it will receive voice messages where user is right now. If normal people don't want this feature they can turn it off. The system work flow is defined in DFD diagrams.

V. DIAGRAMS

A. UML Diagrams:

The Unified Modeling Language (UML) is a general "purpose modeling language in the field of software engineering, which is designed to provide a standard way to visualize the design of a system". It was created and developed by Grady Booch, Ivar Jackson and James Rumbaugh at Rational Software during 1994-1995 with further development led by the through 1996.

VI. RESULTS AND DISCUSSION

- **SPEECH_TO_TEXT Converter:** The system acquires speech at run time through a microphone and processes the sampled speech to recognize the uttered text. The recognized text can be stored in a file. We are developing this on Android platform using Eclipse workbench. Our speech-to-text system directly acquires and converts speech to text. It can supplement other larger systems, giving users a different choice for data entry. A speech-to-text system can also improve system accessibility by providing data entry options for blind, deaf, or physically handicapped users. Speech recognition system can be divided into several blocks: feature extraction, acoustic models database which is built based on the training data, dictionary, language model and the speech recognition algorithm. Analog speech signal must first be sampled at time and amplitude axes, or digitized. Samples of the speech signal are analyzed in even intervals. This period is usually 20 ms because the signal in this interval is considered stationary. Speech feature extraction involves the formation of equally spaced discrete vectors of speech characteristics. Feature vectors from training database are used to estimate the parameters of acoustic models. The acoustic model describes properties of the basic elements that can be recognized. The basic element can be a phoneme for continuous speech or word for isolated words recognition.

- **TEXT_TO_SPEECH Converter:** Converting text to voice output using speech synthesis techniques. Although initially used by the blind to listen to written material, it is now used extensively to convey financial data, e-mail messages, and other information via telephone for everyone. Text-to-speech is also used on handheld devices such as portable GPS units to announce street names when giving directions. Our Text-to-Speech Converterl accepts a string of 50 characters of text (alphabets and/or numbers) as input. In this, we have interfaced the keyboard with the controller and defined all the alphabets as well as digits keys on it. The speech processor has an unlimited Jayachandran .K, Anbumani .P, International Journal of Advance Research, Ideas and Innovations in Technology. © 2017, www.IJARIIT.com All Rights Reserved Page | 1067 dictionary and can speak out almost any text provided at the input most of the times. Hence, it has an accuracy of above 90%. It is a micro controller based hardware coded in Embedded C language. Further research is to be done to optimize various methods of inputting the text i.e. Reading the text using optical sensor and converting it to speech so that almost all sorts of physical challenges faced by the people while communicating are overcome.
- **WORD RECOGNITION:** Voice recognition software (also known as speech to text software) allows an individual to use their voice instead of typing on a keyboard. Voice recognition may be used to dictate text into the computer or to give commands to the computer. Voice recognition software allows for a quick method of writing onto a computer. It is also useful for people with disabilities who find it difficult to use the keyboard. This software can also assist those who have difficulty with transferring ideas onto paper as it helps take the focus out of the mechanics of writing. Word recognition is measured as a matter of speed, such that a word with a high level of recognition is read faster than a novel one. This manner of testing suggests that comprehension of the meaning of the words being read is not required, but rather the ability to recognize them in a way that allows proper pronunciation. Therefore, context is unimportant, and word recognition is often assessed with words presented in isolation in formats such as flash cards Nevertheless, ease in word recognition, as in fluency, enables proficiency that fosters comprehension of the text being read.

VII. CONCLUSION

This paper proposes a system that will be beneficial for society by allowing disable people also to grow along with society. This project makes visually challenged people able enough to be part of growing digital India by allowing them to communicate via internet and also making life of such people much easier. This system overcomes many drawbacks that were faced by visually challenged people such as sending and receiving emails. Success of this project can make an impact on developers motivating them to make something useful that can help visually challenged or blind people.

ACKNOWLEDGEMENT

The authors would like to thank our colleague for implementing the voice based Email for Visually Challenged used in this study. Thanks also to Manish Kumar for helpful discussions and support during this work.

REFERENCES

- [1] G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15– 64.
- [2] The Radicati website Email Statistics Report, 2014-2018 Available:<http://www.radicati.com/wp/wpcontent/uploads/2014/01/EmailStatisticsReport20142018-Executive-Summary.pdf>.
- [3] Ingle, Pranjali, Harshada Kanade, and Arti Lanke. "Voice based e-mail System for Blinds." *International Journal of Research Studies in Computer Science and Engineering (IJRSCE)* (2016): 25-30.
- [4] Isewon, Itunuoluwa, O. J. Oyelade, and O. O. Oladipupo. "Design and implementation of text to speech conversion for visually impaired people." *International Journal of Applied Information Systems* 7.2 (2012): 26-30.
- [5] Shakhovska, N., O. Basystiuk, and K. Shakhovska. "Development of the speech-to-text chatbot interface based on Google API." *CEUR Workshop Proceedings*. Vol. 2386. 2019.
- [6] https://www.juniper.net/documentation/en_US/junos/topics/concept/ipsecauthentication-solutions.html.
- [7] Arlinghaus, Robert, et al. "Understanding the complexity of catch-and-release in recreational fishing: an integrative synthesis of global knowledge from historical, ethical, social, and biological perspectives." *Reviews in Fisheries Science* 15.1-2 (2007): 75-167.
- [8] Cole, Ron, et al. "The challenge of spoken language systems: Research directions for the nineties." *IEEE transactions on Speech and Audio processing* 3.1 (1995): 1-21.
- [9] Lazar, Jonathan, et al. "What frustrates screen reader users on the web: A study of 100 blind users." *International Journal of human-computer interaction* 22.3 (2007): 247-269.
- [10] December, John. "Units of analysis for Internet communication." *Journal of ComputerMediated Communication* 1.4 (1996): JCMC143.
- [11] Pathan, Naziya, et al. "V-Mail (Voice Based E-Mail Application)." (2019).
- [12] Symons, Gary Mark, and Kirk David Symons. "Digital media editing interface using a supercursor for selecting media clips for editing." U.S. Patent No. 8,527,879. 3 Sep. 2013.
- [13] Shabana, T., et al. "Voice based email system for blinds." *International Journal of Advance Foundation And Research In Science & Engineering (IJAFRSE)* Volume 1 (2015).
- [14] Al Smadi, Takialddin, et al. "Artificial intelligence for speech recognition based on neural networks." *Journal of Signal and Information Processing* 6.02 (2015): 66.
- [15] Senders, Joekey T., et al. "Natural and artificial intelligence in neurosurgery: a systematic review." *Neurosurgery* 83.2 (2018): 181-192.
- [16] Collobert, Ronan, et al. "Natural language processing (almost) from scratch." *Journal of machine learning research* 12.Aug (2011): 2493-2537.